

On The Error Propagation in Adams' Extrapolation Method

By B. Zondek and J. W. Sheldon

The error propagation in step by step integration methods is governed by the zeros of certain polynomials [1]. These zeros, $\lambda_0, \dots, \lambda_n$, are the error multipliers, i.e. an initial error propagates through the integration like

$$(1) \quad \eta_r = a_0 \lambda_0^r + a_1 \lambda_1^r + \dots + a_n \lambda_n^r,$$

where r is the step number and a_0, \dots, a_n are constants. In this theory we regard the differential equation as being locally linear with constant coefficient. The Adams extrapolation method [2] of order $n + 1$ for a differential equation

$$(2) \quad y' = f(x, y)$$

is written symbolically

$$(3) \quad y_{r+1} = y_r + h (\beta_0 + \beta_1 \nabla + \dots + \beta_n \nabla^n) f_r,$$

where h is the step size. This leads to the equation

$$(4) \quad \lambda - 1 = k \left\{ \beta_0 + \beta_1 \left(1 - \frac{1}{\lambda}\right) + \dots + \beta_n \left(1 - \frac{1}{\lambda}\right)^n \right\}$$

for the error multipliers $\lambda_0, \dots, \lambda_n$;

$$(5) \quad k = h \frac{\partial f}{\partial y}$$

is considered constant in our approximation. The coefficients β_0, β_1, \dots are generated by the power series expansion

$$(6) \quad \frac{-z}{(1-z) \log(1-z)} = \beta_0 + \beta_1 z + \dots$$

and are all positive. The first few are

$$(7) \quad \beta_0 = 1, \quad \beta_1 = \frac{1}{2}, \quad \beta_2 = \frac{5}{12}, \quad \beta_3 = \frac{3}{8}, \quad \beta_4 = \frac{25}{720}.$$

For sufficiently small $|k|$ (or $|h|$) one of the multipliers, say λ_0 , has a larger absolute value than all the others, $\lambda_1, \dots, \lambda_n$, [3], and an initial error propagates about like $a_0 \lambda_0^r$. In fact it can be shown that for small $|k|$ one root of equation (4) is

$$(8) \quad \lambda_0 = e^k + O(k^{n+2}),$$

while the others are

$$(9) \quad \lambda_r = O((k\beta_n)^{1/n}), \quad r = 1, \dots, n.$$

Consequently, λ_0 causes an error propagation about like that of the differential

equation (2) and we may estimate the accumulated error from the variational differential equation and the local errors.

When we increase $|h|$, keeping n fixed, the multiplier λ_0 differs more and more from e^k . (This has been considered in detail for the Runge-Kutta method in [4], where the notion of fidelity is used.) Furthermore, the "small" multipliers, $\lambda_1, \dots, \lambda_n$, may be expected to increase according to equation (9) and perhaps dominate the "natural" multiplier, λ_0 . In the latter case there is danger of extraneous oscillations appearing in the integration.

When we increase n , keeping h fixed, the multiplier λ_0 approaches e^k according to equation (8) (provided $|h|$ is not too large). On the other hand, from equation (9) and the fact that $\lim (\beta_n)^{1/n} = 1$, we may expect the "small" multipliers to become more significant and perhaps to cause extraneous oscillations.

We now state a lemma that will give us some information about the multipliers when $|k|$ is not small.

LEMMA. Let

- (i) $h(\lambda) = \lambda - 1 - k \left\{ 1 + a_1 \left(1 - \frac{1}{\lambda} \right) + \dots + a_n \left(1 - \frac{1}{\lambda} \right)^n \right\}$,
- (ii) k be real,
- (iii) $k \neq 0$,
- (iv) $k \geq -1$,
- (v) $|k| \{ 2|a_1| + 2^2|a_2| + \dots + 2^n|a_n| \} < 2 + k$;

then $h(\lambda)$ has either one simple zero or no zero outside the unit circle, according as $k > 0$ or $k < 0$ respectively.

PROOF. Let $f(\lambda) = \lambda - 1 - k$ and $g(\lambda) = a_1(1 - 1/\lambda) + \dots + a_n(1 - 1/\lambda)^n$. Thus $h(\lambda) = f(\lambda) - kg(\lambda)$. Now let R be a real number $R > 1$. Then, for $|\lambda| = R$, $\text{Min } |f(\lambda)| \geq |R - 1 - k|$, and this increases without bounds in R . Furthermore, for $|\lambda| = R$,

$$\begin{aligned} \text{Max } |kg(\lambda)| &\leq |k| \left\{ |a_1| \left(1 + \frac{1}{R} \right) + \dots + |a_n| \left(1 + \frac{1}{R} \right)^n \right\} \\ &\leq |k| \{ |a_1| 2 + \dots + |a_n| 2^n \}, \end{aligned}$$

and this is bounded in R . Therefore there exists a real number $R_1 > 1$ such that $|f(\lambda)| > |kg(\lambda)|$ for $|\lambda| \geq R_1$. It is clear from this that $h(\lambda)$ has no zeros on and outside the circle $|\lambda| = R_1$. To prove the lemma, it thus remains to show that $f(\lambda)$ and $h(\lambda)$ have the same number of zeros in the annulus $1 \leq \lambda \leq R_1$. Now by Rouché's theorem [5] (applied here to a doubly connected domain), if $|f(\lambda)| > |kg(\lambda)|$ on the boundary, $f(\lambda)$ and $h(\lambda)$ have the same number of zeros in the domain. The boundary consists of the two circles $|\lambda| = R_1$ and $|\lambda| = 1$. Now R_1 was chosen so that $|f(\lambda)| > |kg(\lambda)|$ on $|\lambda| = R_1$. So it remains to demonstrate this inequality only on the circle $|\lambda| = 1$.

On the unit circle

$$\begin{aligned} \left| \frac{kg}{f} \right| &= |k| \frac{|a_1(1 - 1/\lambda) + \dots + a_n(1 - 1/\lambda)^n|}{|\lambda - 1 - k|} \\ &\leq |k| \frac{|a_1| |\lambda - 1| + \dots + |a_n| |\lambda - 1|^n}{|\lambda - 1 - k|} \equiv \delta. \end{aligned}$$

Now write $r_1 = |\lambda - 1|$ and $r_2 = |\lambda - 1 - k|$. Then it is easily shown that for λ on the unit circle and k real $r_2 = \{(1+k)r_1^2 + k^2\}^{1/2}$, and therefore

$$\left| \frac{kg}{f} \right| \leq \delta(r_1) = |k| \frac{|a_1| r_1 + \cdots + |a_n| r_1^n}{\{(1+k)r_1^2 + k^2\}^{1/2}},$$

where r_1 ranges from 0 to 2. Let us find the maximum of $\delta(r_1)$ in this range. Differentiating we get

$$\frac{d}{dr_1} \delta(r_1) = |k| \frac{k^2 |a_1| + 2k^2 |a_2| r_1 + 3k^2 |a_3| r_1^2 + (4k^2 |a_4| + (1+k) |a_2|) r_1^3 + \cdots + (n-1)(1+k) |a_n| r_1^{n+1}}{\{(1+k)r_1^2 + k^2\}^{3/2}},$$

and, since $k \neq 0$ and $k \geq -1$ by conditions (iii), (iv) of the lemma, $(d/dr_1)\delta(r_1) > 0$, and therefore $\delta(r_1)$, being continuous in the range, has its maximum at $r_1 = 2$. This maximum is

$$\delta(2) = |k| \frac{|a_1| 2 + |a_2| 2^2 + \cdots + |a_n| 2^n}{2+k},$$

and by condition (v) of the lemma $\delta(2) < 1$, so that on the unit circle $|kg/f| \leq \delta(2) < 1$, and the lemma is proved.

Now we apply this lemma to the error propagation in Adams' extrapolation method. There are two cases, $k > 0$ and $k < 0$.

If $k > 0$, we are integrating in the direction of natural error increase ($e^k > 1$). Condition (v) of the lemma becomes

$$(10) \quad k(2\beta_2 + \cdots + 2^{n-1}\beta_n) < 1.$$

This is a sufficient condition that there be only one multiplier outside the unit circle (and this must be the natural one). Consequently, if $|k|^{n+2} \ll e^k$ (see equation (8)), the error propagation is approximately governed by the solution of the variational differential equation. We see, for example, that the 2nd order method ($n = 1$) always satisfies condition (10).

If $k < 0$, we are integrating in the direction of natural error decrease ($e^k < 1$). Condition (v) of the lemma becomes

$$(11) \quad -k(1 + 2\beta_1 + \cdots + 2^n\beta_n) < 2$$

This is a sufficient condition for all multipliers to be within the unit circle. Moreover if k has its critical value

$$(12) \quad k = \frac{-2}{1 + 2\beta_1 + \cdots + 2^n\beta_n}$$

there is a multiplier -1 . (This is seen by substitution into equation (4).) When k is smaller than its critical value, we may expect extraneous oscillations with increasing amplitude. Accordingly, if we want the error propagation to be approximately governed by the variational differential equation, it is certainly necessary for condition (11) to be satisfied (in addition, of course, to the fidelity condition

$|k|^{n+2} \ll e^k$.) We get, for example, for the 4th order method ($n = 3$) the critical value $k = -\frac{3}{10}$. We also note that the critical value of k tends to zero as the order is increased because $\sum_r 2^r \beta_r$ is a divergent series.

In practice, for n not too small, condition (11) is almost necessary and sufficient for the natural multiplier, $\lambda_0 = e^k + 0(k^{n+2})$, to dominate the other multipliers.

When $f(x, y)$ is complicated, the labor of integration is proportional to $1/h$. If we minimize this, subject to given truncation error and one of the conditions (10) or (11), we get optimum values of h and n . This type of optimization is discussed for orbit integrations in [6].

Naval Proving Ground, Dahlgren, Virginia and
Computer Usage Company, New York, New York

1. L. COLLATZ, *Numerische Behandlung von Differentialgleichungen*, 2nd edition, Springer, Berlin, 1955, p. 106 and p. 128. References to several papers on the subject are given there.
2. L. COLLATZ, *Numerische Behandlung von Differentialgleichungen*, 2nd edition, Springer, Berlin, 1955, p. 79.
3. H. RUTISHAUSER, "Ueber Instabilitaet von Methoden zur Integration gewoehnlicher Differentialgleichungen," *Zeit. f. ang. Math. u. Phys.*, v. 3, 1952, p. 65-72.
4. HARRY L. REED, *Numerical Integration of Oscillatory Systems*, Ballistic Research Laboratory Report No. 957, October 1955.
5. Z. NEHARI, *Conformal Mapping*, McGraw-Hill Book Co., Inc., New York, 1952, p. 131.
6. J. W. SHELDON, B. ZONDEK & M. FRIEDMAN, "On the time step to be used for computation of orbits by numerical integration," *MTAC*, v. XI, 1957, p. 181-189.